

C 4

Catalogue 92-950



Mother Tongue

1991 Census Technical Reports



Census



Statistics
Canada

Statistique
Canada

Canada

How to Reach Us

Data in Many Forms...

Statistics Canada disseminates data in a variety of forms. In addition to diskette products and publications, both standard and special tabulations are available in print or electronic format. Maps and other geographic reference materials are available for some types of data. In some cases, workshops and seminars have been developed to provide additional information on some subjects.

How to Obtain More Information

Inquiries about products and services should be directed to your local Statistics Canada Reference Centre:

St. John's	1-800-565-7192	Winnipeg	(204) 983-4020
Halifax	(902) 426-5331	Regina	(306) 780-5405
Montréal	(514) 283-5725	Edmonton	(403) 495-3027
Ottawa	(613) 951-8116	Calgary	(403) 292-6717
Toronto	(416) 973-6586	Vancouver	(604) 666-3691

For users who reside outside the local dialing area of the Regional Reference Centres, a toll-free number is provided:

Newfoundland, Labrador, Nova Scotia, New Brunswick and Prince Edward Island	1-800-565-7192
Quebec	1-800-361-2831
Ontario	1-800-263-1136
Manitoba, Saskatchewan, Alberta and Northwest Territories	1-800-563-7828
British Columbia and Yukon	1-800-663-1551
Telecommunications Device for the Hearing Impaired	1-800-363-7629
Toll Free Order Only Line (Canada and United States)	1-800-267-6677

How to Order Products

Statistics Canada publications may be purchased from local authorized agents and other community bookstores, through the nearest Statistics Canada Regional Reference Centre, or by mail order to Marketing Division, Sales and Service, Statistics Canada, Ottawa K1A 0T6. Contact the nearest Statistics Canada Regional Reference Centre for more information on electronic products and services. See the order form at the end of this product for further details.



Statistics Canada

1991 Census Technical Reports

Mother Tongue

Reference Products series

Published by authority of the Minister responsible for Statistics Canada

©Minister of Industry, Science and Technology, 1993

All rights reserved. No part of this document may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission from Licence Services, Marketing Division, Statistics Canada, Ottawa, Ontario, Canada K1A 0T6.

August 1993

Price: Canada: \$20.00
United States: US\$24.00
Other Countries: US\$28.00

Catalogue 92-335E

ISBN 0-660-14268-6

Ottawa

La version française de cette publication est disponible sur demande (ISBN 0-660-93518-X)

Note of Appreciation

Canada owes the success of its statistical system to a long-standing cooperation involving Statistics Canada, the population of Canada, its businesses and governments. Accurate and timely statistical information could not be produced without their continued cooperation and goodwill.

Canadian Cataloguing in Publication
Data

Main entry under title:

Mother tongue

(Reference product series)
(1991 census technical report)
Issued also in French under title: Langue
maternelle.
ISBN 0-660-14268-6
CS92-335E

1. Native language -- Statistics. 2. Canada -- Languages -- Statistics. 3. Language surveys -- Canada.
4. Canada -- Census, 1991. I. Statistics Canada.
- II. Series. III. Series.

HA741.5.1991 M67 1993 305.7'0971'021
C93-099359-4

How to Cite This Document

Statistics Canada. Mother Tongue. 1991 Census Technical Reports;
Reference Products Series. Ottawa: Minister of Industry,
Science and Technology, 1993. Catalogue number 92-335E.

The paper used in this user documentation meets the minimum requirements of American National Standard for Information Sciences - Permanence of Paper for Printed Library Materials. ANSI Z39.48-1984



Preface

Through time, the Census of Canada has become the primary source of information about Canadians and how they live. Decisions based on this information affect the social and economic affairs of all Canadians.

Statistics Canada, as the professional agency in charge of producing this information, has the responsibility for informing users of data quality. The agency must describe the concepts and methodology used in collecting and processing the data, as well as any other features that may affect their use or interpretation.

In order to describe the quality of the 1991 Census data, Statistics Canada has prepared the following publications: a census **Dictionary**, which provides concise and easy to understand textual and graphical information pertaining to census concepts; a **Handbook**, which provides an overview of how the census is conducted; and a series of **Technical Reports**, which present in greater detail, information on the quality of data for specific characteristics, such as mother tongue, as covered in this report.

Information on data quality is important for users. It allows them to assess the usefulness of census data for their purposes as well as the risks involved in basing conclusions or decisions on these data. The 1991 Census was a large and complex undertaking and, while considerable effort was taken to ensure high standards throughout all collection and processing operations, the resulting data are inevitably subject to a certain degree of error.

Information on data quality is also important to Statistics Canada. It is an integral part in the development and maintenance of pertinent and reliable statistical programs.

This publication is a major contribution to achieving these goals. It has been prepared by Raj K. Chawla of the Demolinguistics Division and Dr. John de Vries of Carleton University. The authors would like to acknowledge the contribution of Mr. Réjean Lachapelle, Director, Demolinguistics Division. Also, the authors wish to thank Pierre Turcotte for compiling the initial data, and the support staff from three Divisions in Statistics Canada: Demolinguistics, Census Operations and Social Survey Methods.

Finally, I would like to express my appreciation to the millions of Canadians who completed their questionnaires on June 4, 1991, as well as to those who assisted Statistics Canada in planning and conducting the census.

Ivan P. Fellegi
Chief Statistician of Canada

Table of Contents

	Page
List of Tables	(ii)
List of Figures	(ii)
I. Introduction	1
II. Concepts and Definitions	3
III. Data Collection and Coverage	7
IV. Data Assimilation	11
V. Edit and Imputation	14
VI. Data Evaluation	18
VII. Historical Comparability	25
VIII. Products and Services	30
IX. Conclusion	31
Bibliography	32
Other Census Reference Products	33
Regional Reference Centres	35
Depository Libraries	36

List of Tables

	Page
Tables	
2.1 Language Questions in the 1986 and 1991 Censuses	6
3.1 Distribution of Non-permanent Residents by Mother Tongue, Canada, Provinces and Territories, 1991 Census	9
3.2 Estimated Population Undercoverage Rates by Mother Tongue, Canada, 1981, 1986 and 1991 Censuses	10
4.1 Overview of Automated Coding on Mother Tongue Variable, 1991 Census	13
5.1 Imputation for Mother Tongue by Type, Canada, Quebec and Canada Excluding Quebec, 1991 Census	16
5.2 Imputation for Mother Tongue by Type, for Persons With Single Response, 1991 Census	16
5.3 Imputation for Mother Tongue by Type and Age Groups, 1991 Census	17
6.1 Non-response Rates for Mother Tongue, Canada, Provinces and Territories, 1981, 1986 and 1991 Censuses	19
6.2 Multiple Response Rates for Mother Tongue, Canada, Provinces and Territories, 1981, 1986 and 1991 Censuses	20
6.3 Distribution by Type of Invalid Responses to Mother Tongue, 1991 Census	21
6.4 Distribution of Write-ins of English and/or French as Mother Tongue, as Captured and After Corrections, 1991 Census	21
6.5 Distribution of Population by Mother Tongue, Before and After Imputation, 1991 Census (100% Data)	22
6.6 Distribution of Population by Categories of Mother Tongue and Sources of Data, Canada, Provinces and Territories, 1991 Census	24
7.1 Prevalence of Multiple Responses to Mother Tongue Questions by Form, Canada, 1981, 1986 and 1991 Censuses	26
7.2 Distribution of Population by Mother Tongue, 1981, 1986 and 1991 Censuses	29

List of Figures

Figures

1. Schematic Process of Family Imputation	15
2. Comparison of Non-response Rates for Mother Tongue	18
3. Distribution of Population by Mother Tongue	28

I. Introduction

Since 1971, Statistics Canada has provided the public with various forms of census data quality reports. Over the years, the reliance on these reports as data quality reference tools has become paramount.

The Canadian Census of Population is a major undertaking consisting of several phases, each one of which is complex and consumes a significant amount of time and resources. The desired information is sought through a set of questions established after detailed consultations and tests; is collected in a specific manner through trained enumerators; is checked for inconsistencies and errors, with the final results placed on a computer database and selected data analyzed; and published in a series of reports.

The primary objectives of the 1991 Census were to obtain accurate counts of the population, and to determine the number of households and dwellings at all geographic levels, as well as to collect a broad range of information on their characteristics. The census is an invaluable source of information that is useful to the various levels of government, to business, associations, interest groups, and the general public. The data can be used in government planning of social and economic programs, assessment of the need for educational and health facilities, and planning by private enterprise.

The national census provides the most comprehensive database on the characteristics of Canadians, their families and their households. The information ranges from age and sex of individuals to their ethnic origin, education, occupation, labour force activity, industry, sources of income, and includes their family and household characteristics.

However, in a massive project such as the census, the results are never perfect. Although considerable effort has been made to maintain high standards of quality, errors inevitably occur at various stages of the collection and processing operations. Users should be aware of the nature and scope of any errors that the census data may contain, as well as the risks involved in basing conclusions or decisions on these data.

In order to inform data users of the potential problems or intricacies of the data, a number of programs for assessing the quality of census data have been developed.

The **1991 Census Technical Reports** inform users of the conceptual framework and definitions used in the data collection.

Users are generally aware that data are subject to error; documentation is provided in these reports about the overall accuracy of the data. Likely principal sources of error are indicated and, where possible, the potential size of the error is given. Furthermore, any unusual circumstances which might influence the data are identified.

Users are informed of the data collection and processing methodology so that they can verify whether the data adequately approximate what they wish to measure and whether the estimates they wish to use were produced with tolerances acceptable for their intended purposes.

In this technical report, we will discuss various aspects of the quality of data on **mother tongue**. The question on **mother tongue** was included in both the short (2A) and long (2B) questionnaires used in the 1991 Census. Besides the question on **mother tongue**, the 2B questionnaire included three other questions on language, namely on language spoken most often at home, on knowledge of English and French, and on knowledge of language(s) other than English or French. These four language questions, complemented by questions on ethnicity, religious affiliation and immigration provide an opportunity to study linguistic and cultural characteristics of Canadians. The four questions on languages are designed to collect the **demolinguistic** data. Demolinguistics, a subdiscipline of demography (not of linguistics), involves the demographic analysis of data on languages. Such analysis is useful for our understanding of, for instance, the linguistic diversity of Canadians, the evolution of language groups, or the transmission of mother tongue between generations.

Census of Population - Reference Products
1991 Census Technical Reports

The report follows, in general, the stages of data collection and assimilation as they are used in the census procedures. The first section gives a description of the essential concepts and their definitions. This is followed by a brief description of the procedures of data collection and some aspects of **coverage**, in particular those which may affect the data on **mother tongue**. A brief description of the processing stages of the data assimilation operation, in regional offices and at Statistics Canada's head office, again emphasizes the impact of these processes on the **mother tongue** data. A separate section discusses the procedures used for editing and imputing **mother tongue** data. The final sections of this report deal with overall quality and historical comparability of the data. In particular, the effects of changes and improvements in data collection methods on the comparability of the 1991 Census data with those from preceding censuses will be discussed.

II. Concepts and Definitions

1. Demolinguistic concepts

Demolinguistic data may be classified into three categories:

- (i) mother tongue, pertaining to language first learned in early childhood;
- (ii) usual language, referring to **current** language use of the respondent;
- (iii) ability to speak one or more designated languages.

The United Nations' recommendations on demolinguistics data, made with regard to censuses to be taken in 1970, have not been altered substantially in subsequent publications:

"224. There are three types of language data which can be collected in censuses. These are:

- (a) mother tongue, defined as the language usually spoken in the individual's home in his early childhood;
- (b) usual language, defined as the language currently spoken, or most often spoken, by the individual in his present home; and
- (c) ability to speak one or more designated languages.

"225. Each of these types of information serves a distinctly different analytical purpose. Each country should decide which, if any, of these types of information is applicable to its own needs. International comparability of tabulations is not a major factor in determining the form of the data to be collected on this topic.

"226. In the compilation of data on usual language or on mother tongue, it is desirable to show each language that is numerically important in the country and not merely the dominant language.

"227. Information on language should be collected for all persons. In the tabulated results for children under five years of age, the criterion for determining language for children not yet able to speak should be clearly indicated." (United Nations, 1969, p. 21.)

It should be noted that Canadian census practices have followed these recommendations very closely. Canada is the only country in which census data are collected on all three aspects. Data on mother tongue and on language spoken most often at home have been collected and reported for over one hundred categories of languages.

2. Definitions

In the 1991 Census, **mother tongue** refers to the first language learned at home and still understood by the individual at the time of the census. As per instructions in the 1991 Census Guide, a respondent must report the first language learned at home before starting school. If this language was no longer understood, he/she must report the second language learned. If two languages were learned at the same time and were used equally often, he/she could report both (for more details, see **Mother Tongue, Catalogue No. 93-313**).

The census definition of mother tongue, therefore, pertains to the early childhood of the person concerned. The mother tongue is the language transmitted at home by parents and other family members to an individual during childhood. In cases in which more than one language is spoken regularly in a family, the definition requires clarification. Emphasis may be put on either the criterion of precedence (i.e. the **first** language learned) or on the criterion of pre-eminence (that is, the language spoken **most often** during childhood). In the latter case, a choice must be made between two referents: the individual or the entire family. The United Nations recommendations prefer the latter. The definition it provides has, however, one major drawback: in immigrant families, it is not

uncommon for the parents to communicate with each other in a language which they neither speak to, nor teach, their children. As adults, the children of such families may then report as their mother tongue a language which they never really spoke (Lachapelle, 1991).

Finally, individuals may be classified in two ways with regard to mother tongue:

- the language first learned or spoken at home during early childhood (criterion of precedence); or
- the language spoken most often at home in early childhood (criterion of pre-eminence).

The former approach is probably more generous to proportionately small minority language communities. It is, however, the case that any definition of mother tongue will have to allow for the possibility that some persons have more than one mother tongue; this will be especially the case in families in which the two parents have different mother tongues and used different languages in the home.

Canadian censuses have used definitions based on the criterion of precedence. The following is the question asked in the 1991 Census:

"What is the language that this person **first learned** at home in **childhood** and **still understands**?"

If this person no longer understands the first language learned, indicate the second language learned.

This phrasing is similar to that used in the censuses of 1981 and 1986, except for the expression "at home". The wording of the question has not changed very much since 1941. The censuses of 1941, 1961, 1981 and 1986 referred to the language **first learned**, whereas those of 1951, 1971 and 1976 referred to the language **first spoken**. In all cases, the condition that the language be **still understood** was included. This clause distinguishes the Canadian mother tongue data from those obtained in other countries (and, of course, from the phrasing recommended by the United Nations).

The inclusion of the "still understood" clause may be explained on the grounds of retaining historical consistency in the question's wording through the early 1960s, and on the basis of Section 23 of the **Canadian Charter of Rights and Freedom** (1982). This section refers to the "first language learned and still understood" in relation to minority language educational rights. The phrasing thus defined does allow for "mother tongue loss" in cases in which individuals no longer understand the language they first learned or spoke in early childhood. Lachapelle (1991, pp. 35-38) has shown, on the basis of more detailed language data from the 1986 General Social Survey, that such mother tongue loss is overall an uncommon phenomenon, but that it does have a slight effect on French mother tongue minorities in regions in which they represent less than 5% of the total population.

The 1991 Census question on mother tongue differed from that asked in preceding censuses in several ways:

- (i) Beginning with the census of 1971, the question on mother tongue was the only language question which was asked of all respondents; the other demolinguistic questions were asked (with the exception of the 1976 Census in which only mother tongue was asked) to respondents in a sample of one-third of households in 1971 and one-fifth of households in the 1981, 1986 and 1991 Censuses. In 1991, the ordering of questions on languages in the 2B questionnaire changed in such a way that the question on mother tongue was preceded by questions on the knowledge of English or French, knowledge of language(s) other than English or French, and language spoken most often at home. The four questions on languages were placed together as a **block** in order to help respondents provide precise information on the knowledge and use of different languages. The objective was to improve data quality and to minimize the proportion of respondents entering multiple responses. The objective was accomplished; for instance, the proportion of the population reporting more than one mother tongue (in other words, giving multiple responses to the question on mother tongue) fell to 1.2% in 1991 from 3.4% in 1986 (see Section VI of this report). However, studies have shown that such multiple responses were highly unreliable (Lachapelle, 1991).
- (ii) In the 2B questionnaire of the 1991 Census, the block containing questions on languages (including mother tongue) was titled (in bold capitals) **LANGUAGE** whereas in the 2A questionnaire, the question on mother tongue carried the caption (in bold capitals) **LANGUAGE FIRST LEARNED AT HOME IN CHILDHOOD**. No such titles/captions were used in the 1986 Census, as shown in Table 2.1.

Overall, two important changes were introduced in the 1991 Census:

- (i) a new question dealing with the respondent's ability to speak non-official languages (parallel to the question on the ability to speak the official languages) was introduced;
- (ii) in the "long form" (2B) questionnaire, a module of language questions was created in order to improve the quality of the responses. This module consisted of the following questions:
 - the question on the ability to speak the official languages;
 - the question on the ability to speak other non-official languages;
 - the question on the language spoken most often at home;
 - the question on mother tongue.

One effect of the change in format of the question on **mother tongue** between the 1986 and 1991 Censuses is that the data from the 2B questionnaire in 1991 are not strictly comparable with the data from earlier censuses. Data based on the 2A questionnaire, on the other hand, are more comparable. This point becomes especially important for analyses involving the joint use of mother tongue data and other demolinguistic variables since such analyses obviously have to use data compiled from the 2B questionnaire. The reasons underlying differences in estimates compiled from 2A and 2B questionnaires are highlighted in Section VI and the problem of comparing the 1991 estimates on mother tongue with those from earlier censuses is discussed in Section VII of this report.

In the 1991 Census, instructions to the respondent specified that "if this person no longer understands the first language learned, indicate the second language learned". Additional commentary included in the census guide which accompanied the questionnaire provides the further specification that "childhood" refers to the years before the respondent started attending school. Furthermore, detailed instructions specify under what conditions more than one mother tongue should be reported (only if two languages were learned at the same time and were used equally often at home by the respondent in early childhood).

The guide gave instructions regarding the data to be supplied for children who have not yet learned a language: the first language such a child will learn **at home** should be mentioned as the child's mother tongue. Here, too, more than one mother tongue should be reported only if the child will learn two languages at the same time and if these languages will be spoken equally often.

Similarly, if a person spoke an **Amerindian** (North American Indian) language, he/she was to report the specific Amerindian language such as Cree or Ojibway. Again, persons from India were not to report **Indian** as a language but rather a specific language such as Hindi, Urdu or Punjabi.

Table 2.1 Language Questions in the 1986 and 1991 Censuses

Questionnaire	1986	1991
2A	<p>6. What is the language you first learned in childhood and still understand? (See <i>Question Guidelines</i>)</p> <ul style="list-style-type: none"> * English * French * Italian * German * Ukrainian * Other (specify) 	<p>LANGUAGE FIRST LEARNED AT HOME IN CHILDHOOD?</p> <p>7. What is the language that this person first learned at home in childhood and still understands? (If this person no longer understands the first language learned, indicate the second language learned.)</p> <ul style="list-style-type: none"> * English * French * Other - <i>Specify</i>
2B (20% Sample)	<p>6. What is the language you first learned in childhood and still understand? (See <i>Guide</i>)</p> <ul style="list-style-type: none"> * English * French * Italian * German * Ukrainian * Other (specify) <p>18. What language do you yourself speak at home now? (If more than one language, which language do you speak most often?) (See <i>Guide</i>)</p> <ul style="list-style-type: none"> * English * French * Italian * Chinese * German * Other (specify) <p>19. Can you speak English or French well enough to conduct a conversation? (See <i>Guide</i>)</p> <p>Mark one box only</p> <ul style="list-style-type: none"> * English only * French only * Both English and French * Neither English nor French 	<p>LANGUAGE</p> <p>7. Can this person speak English or French well enough to conduct a conversation? Mark one circle only.</p> <ul style="list-style-type: none"> * English only * French only * Both English and French * Neither English nor French <p>8. What language(s), other than English or French, can this person speak well enough to conduct a conversation? * None OR * Specify other language(s) (up to three languages)</p> <p>9. What language does this person speak most often at home? * English * French * Other - <i>Specify</i></p> <p>10. What is the language that this person first learned at home in childhood and still understands? (If this person no longer understands the first language learned, indicate the second language learned.)</p> <ul style="list-style-type: none"> * English * French * Other - <i>Specify</i>

III. Data Collection and Coverage

1. Data Collection

In the 1991 Census, information on individuals and households was collected at their usual place of residence - this method of collection is known as the "de jure" approach. As much as possible, data were also collected on residents without a fixed address. For the majority of private households, a self-enumeration method was used. A member of the household was to complete the questionnaire as of June 4th, 1991, and to mail back the completed form in a pre-addressed envelope. In remote or northern areas and on Indian reserves, enumeration was by canvasser. Less than 2% of all households were enumerated by the canvasser method. A special form (Form 3B) was used to enumerate homeless people living, for example, in abandoned buildings or shelters, or using soup kitchens.

The regular collection methods used two census forms:

- the "short form" (Form 2A), which was administered to 80% of all private households and contained questions on name, relationship to a reference person, date of birth, sex, legal marital status, common-law status, mother tongue, household maintainer(s) and dwelling tenure;
- the "long form" (Form 2B), given to 20% of all private households, which contained all of the questions appearing on the Form 2A plus additional questions on labour force activity, income, education, disability, citizenship, housing, ethnic origin, language, etc. (Note that the mother tongue question appeared on both the short form and the long one.)

In addition, special questionnaires were used for the enumeration of Canadians posted abroad in the Armed Forces or in government service (Form 2C) and for the enumeration of non-institutional collective dwellings (such as hotels, rooming-houses and the like).

2. Coverage

An important aspect of any census is the extent to which the collected data refer to the intended population, or its coverage.

In 1991, several initiatives were taken to improve coverage.

These included:

- using paid advertising to inform Canadians on when and how "to count themselves in";
- creating an address register from other sources of information and using this list to check if any dwellings were missed;
- establishing special procedures to count homeless people through soup kitchens;
- establishing special procedures to count the population on Indian reserves;
- using a respondent-friendly questionnaire;
- offering respondents information through a Public Communications Program and a multilingual Telephone Assistance Service.

Despite these and other improvements, it is virtually impossible to achieve 100% coverage in such a large survey. Consequently, various checks are performed on the collection of data, including those required to detect the undercoverage and overcoverage of data. These coverage problems are defined as follows:

- (i) **undercoverage** is the extent to which individuals who should have been included are in fact missing;

Census of Population - Reference Products
1991 Census Technical Reports

(ii) **overcoverage** is the extent to which persons are included who were not part of the intended population, or to which individuals are counted more than once (for example, university students living away from home who were enumerated at their parents' address as well as at their university residence).

Several aspects of coverage should be considered in the use of the mother tongue data for 1991:

(a) On some Indian reserves and Indian settlements, enumeration was not permitted or was interrupted before it could be completed. In total, 78 geographic areas were thus affected. A similar problem arose in the 1986 Census and, to a lesser extent, in 1981 as well. Although the total population in these areas was not large, these are of course areas in which a high proportion of the inhabitants had an aboriginal language as mother tongue. Given that among the enumerated population there were about 200,000 persons with an aboriginal language as mother tongue, undercoverage of these language communities as a consequence of the incomplete enumeration of Indian reserves may have been significant, especially for smaller areas where the affected reserves and settlements accounted for a higher proportion of the population.

(b) A second difference with respect to earlier census practices is the inclusion of both permanent and non-permanent residents (in preceding censuses, only permanent residents were included). Non-permanent residents include persons with student or employment authorizations, holders of a Minister's permit and refugee claimants. According to the 1991 Census, there were 223,410 non-permanent residents who, in turn, represented less than 1% of Canada's total population. Approximately 70% of these persons had a mother tongue other than English or French. Overall, 56.5% of these persons were found to be living in Ontario, 19.7% in Quebec, 12.6% in British Columbia, and the remaining 11% in other parts of Canada (Table 3.1).

In addition to these two special circumstances, censuses are always faced with a certain amount of undercoverage, as well as a much smaller amount of overcoverage. Special studies, such as the Reverse Record Check and the Overcoverage Study, are conducted to produce estimates of undercoverage, overcoverage and net undercoverage (i.e. the difference between the two components).

According to the issue of the Daily dated November 4, 1992, the national rate of net undercoverage in the 1991 Census was 3.2%. This rate, however, varied across provinces - from 1.2% for Prince Edward Island to 4.1% for Ontario. For Canada, excluding the Yukon Territory and the Northwest Territories, the **gross** undercoverage rate (i.e. without including the effects of overcoverage, and excluding non-permanent residents, and treating incompletely enumerated Indian Reserves as missed in both censuses) was 3.45%, virtually unchanged from the 1986 rate of 3.38% (Table 3.2). This nationally very small intercensal change had masked somewhat larger changes at the provincial level (for example, the undercoverage increased significantly in Ontario and decreased significantly in British Columbia).

For the 1991 Census, the gross undercoverage rates for the two largest language groups (namely, English and French) were lower than the national rate. The rate for persons with French as mother tongue was slightly higher than for those with English as mother tongue. The undercoverage rate for persons with mother tongue neither English nor French was 1.5 times the national rate; one of the reasons for a higher rate for this group was that it contained a large proportion of immigrants. Studies of undercoverage have found that recent immigrants are less likely to be enumerated than is the rest of the population (see 1986 User Information Bulletin, Number 2, July 1988).

Table 3.1 Distribution of Non-permanent Residents by Mother Tongue, Canada, Provinces and Territories, 1991 Census(1)

	Single response				Multiple response					Total	
	English	French	Non-official language	Total	English and French	English and non-official language	French and non-official language	English, French and non-official language	Non-official languages	Total	
	Numbers										
Canada	53,045	6,635	156,725	216,410	300	4,715	1,070	180	740	7,000	223,410
Newfoundland	420	60	345	820	0	0	0	5	0	10	830
Prince Edward Island	70	5	50	120	0	0	0	0	0	0	120
Nova Scotia	765	50	805	1,625	0	60	0	0	10	75	1,700
New Brunswick	675	265	405	1,345	5	5	5	0	0	15	1,355
Quebec	5,405	4,860	31,855	42,120	165	625	825	135	95	1,850	43,970
Ontario	32,670	950	88,535	122,155	90	3,230	160	30	495	4,005	126,165
Manitoba	1,175	55	2,720	3,950	10	50	0	0	20	75	4,025
Saskatchewan	1,095	35	1,730	2,860	0	15	10	0	0	20	2,875
Alberta	3,845	115	9,725	13,680	5	315	25	0	50	400	14,065
British Columbia	6,820	225	20,445	27,490	25	405	50	5	65	550	28,040
Yukon Territory	40	20	30	90	0	0	0	0	0	0	90
Northwest Territories	70	5	70	150	0	0	0	0	0	0	150
% Distribution											
Canada	23.7	3.0	70.2	96.9	0.1	2.1	0.5	0.1	0.3	3.1	100.0
Newfoundland	50.6	7.2	41.6	98.8	0.0	0.0	0.0	0.6	0.0	1.2	100.0
Prince Edward Island	58.3	4.2	41.7	100.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
Nova Scotia	45.0	2.9	47.4	95.6	0.0	3.5	0.0	0.0	0.6	4.4	100.0
New Brunswick	49.8	19.6	29.9	99.3	0.4	0.4	0.4	0.0	0.0	1.1	100.0
Quebec	12.3	11.1	72.4	95.8	0.4	1.4	1.9	0.3	0.2	4.2	100.0
Ontario	25.9	0.8	70.2	96.8	0.1	2.6	0.1	0.0	0.4	3.2	100.0
Manitoba	29.2	1.4	67.6	98.1	0.2	1.2	0.0	0.0	0.5	1.9	100.0
Saskatchewan	38.1	1.2	60.2	99.5	0.0	0.5	0.3	0.0	0.0	0.7	100.0
Alberta	27.3	0.8	69.0	97.1	0.0	2.2	0.2	0.0	0.4	2.8	100.0
British Columbia	24.3	0.8	72.9	98.0	0.1	1.4	0.2	0.0	0.2	2.0	100.0
Yukon Territory	44.4	22.2	33.3	100.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
Northwest Territories	46.7	3.3	46.7	100.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

(1) Based on 20% sample data.

Historically, for groups with English and French as mother tongue, the undercoverage rates from the 1981 Census were the lowest compared to their respective counterparts from the 1986 and 1991 Censuses. The rates for English and French categories were not significantly different in the 1991 Census, nor were they in the 1986 and 1981 Censuses. Again, for both 1991 and 1981 Censuses, the rate for persons with neither English nor French as mother tongue was 1.5 times the national rate whereas in the 1986 Census, such a straight comparison was not feasible - because of the difference in the definition of "other" group (see Table 3.2).

Census of Population - Reference Products
1991 Census Technical Reports

Table 3.2 Estimated Population Undercoverage Rates(1) by Mother Tongue, Canada(2), 1981, 1986 and 1991 Censuses

Mother tongue	1991		1986		1981	
	Rate	Standard error	Rate	Standard error	Rate	Standard error
percentage						
Total	3.45	0.13	3.38	0.12	2.01	0.09
English	3.12	0.21	3.12	0.13	1.86	0.11
French	3.20	0.22	3.10	0.33	1.80	0.20
Italian			1.90	0.77		
German			1.15	0.52		
Ukrainian			2.10	0.94		
Other (3)	5.04	0.32	7.62	0.58	3.06	0.26

(1) Estimates are subject to change.

(2) Excluding population of the Territories and non-permanent residents.

(3) Including Italian, German and Ukrainian in 1991 and 1981.

IV. Data Assimilation

Data assimilation is the processing phase during which data from the census questionnaires are edited, coded and captured. The process includes the transformation of the questionnaire responses into machine-readable form.

The five components of data assimilation are:

- Regional Office Processing (ROP);
- Direct Data Entry (DDE);
- Head Office Processing;
- Automated Coding;
- Classification.

1. Regional office processing (ROP)

Regional processing centres handled part of the data assimilation (i.e. the processing of the completed census forms). Data assimilation consisted of the following:

(i) Receipt and Preparation of Documents

When completed questionnaires reached the regional processing centres, they were logged, counted and prepared for key entry. Preparation involved checking questionnaires against Visitation Records – making sure, for example, that the number of household members on the two documents matched. Legibility checks ensured that the data were ready for computer entry. Finally, all responses concerning relationships between household members (Question 2) were converted to numerical codes.

(ii) Reverse Record Check

A sample of persons was taken from the 1986 Census files and other external sources; 1991 documents were then searched for these same persons. Where a person was found, 1991 characteristics were noted and sent to head office. Where persons were not found, further tracing determined whether they were enumerated elsewhere in Canada or missed altogether.

(iii) Coding of Economic Variables

Supervisors and coding consultants resolved any discrepancies in coding before the questionnaires for an enumeration area proceeded to the next stage. Other sources (city directories, subject-matter officers, and so on) were occasionally consulted.

(iv) Processing

Questionnaires were transferred to work units in charge of direct data entry at Revenue Canada Taxation (RCT) regional processing offices, before being sent to Statistics Canada in Ottawa.

2. Direct data entry (DDE)

Direct entry (key entry) of data from the census of population questionnaires was performed by 1,500 operators, sworn to secrecy under the **Statistics Act**, working for Revenue Canada in seven centres across the country. Data were then transmitted from the regional centres to Ottawa in order to be stored on tapes.

The activities that were involved in this stage include:

- receipt and registration of enumeration area boxes from regional processing;
- key entry of census questionnaire data by operators;
- sample verification of the captured data by a different operator, including the correction of flagged errors by the system at operator entry;
- return of questionnaires to enumeration area boxes (in their original order);
- shipping to Ottawa (shipment of data cartridges between Revenue Canada Taxation (RCT) head office in Ottawa and Statistics Canada head office, on a daily basis).

3. Head office processing (HOP)

Head office processing is a combination of automated and manual processing designed to carry out structural edits on the census data and to process special enumeration returns (for example those received from Canadians living abroad, temporary residents, and personnel from merchant or navy ships).

At Statistics Canada's head office in Ottawa, Visitation Records and questionnaires were received, registered and stored. Tape cartridges with data were copied onto the head office processing database. Automated structural edits were carried out at the enumeration area level, as well as at the household and individual level, to identify and resolve inconsistencies.

Head office processing was performed in three phases, namely, "Receipt, Registration and Storage", "Data Analysis" and "Special Processing". Coverage studies were also conducted and necessary adjustments made.

4. Automated coding

An innovation in data assimilation was the Automated Coding by Text Recognition (ACTR) of data on 13 variables. These included data on mother tongue from the 2A and 2B questionnaires, language spoken at home, and knowledge of language(s) other than English or French from the 2B questionnaire. Data on the knowledge of official language(s) were not processed by this method since there were no write-in responses involved.

A major modification for language questions in 1991 was that two write-in responses were key-entered, whereas only one write-in response was processed in the preceding censuses. The words thus entered were **parsed**, i.e. standardized by means of custom rules; extraneous text was removed and, where appropriate, semantically equivalent information was substituted. The result was a set of words which could be matched against a **reference file**, i.e. a list of recognized terms with associated numerical codes (see Ciok, 1991, for a detailed description). The coding procedure used a **matching algorithm** to determine the degree of correspondence between a parsed root and the entries on the reference file, and a facility for **code assignment**.

All responses which could not be coded by the system were coded manually by processing staff and coding consultants using a computer-assisted process. Once all responses were coded, an analysis was conducted for each variable to ensure compliance with rules governing data quality, and then records were forwarded to Edit and Imputation.

The precoded mother tongue categories (English and French) did not require further coding. However, automated coding was applied to all of the write-in responses to the mother tongue question, on Forms 2A and 2B. In total, 3,175,860 write-in responses were coded for 2A and 1,010,324 for 2B (Table 4.1).

Table 4.1 Overview of Automated Coding on Mother Tongue Variable, 1991 Census

Questionnaire	Responses coded		
	Autocoding	Manual	Total
2A			
Number	2,998,446	177,414	3,175,860
Distribution (%)	94.4	5.6	100
2B			
Number	958,669	51,655	1,010,324
Distribution (%)	94.9	5.1	100

Source: 1991 Automated Coding - MIS Executive Summary, Week Ending February 7, 1992.

It may be noted from Table 4.1 that the write-in responses in respect to **mother tongue** in both questionnaires 2A and 2B, which were coded by the automated system, represented between 94% and 95% of the total write-in responses; the remaining 5% to 6% write-ins were manually coded. The combined system and manual error rate in respect to coding of the mother tongue variable was far less than one percent.

It is hard to assess the impact of these procedures on the quality of the mother tongue data. An initial assessment suggests that the drastic reduction in human intervention in the coding process produced a higher reliability for these data than was achieved in the earlier censuses (using manual coding procedures). However, without systematic comparisons of the two approaches, it is not possible to give a more accurate evaluation. It should be pointed out though that the heavy use of the automated coding procedure (which relies on predesigned specifications) is more likely to produce systematic errors than random errors.

5. Classification

In comparison with the procedures used in the 1986 Census, there were some changes in the way in which languages were categorized. In both the 1986 and 1991 Censuses, languages were grouped by language family (Romance languages, Dravidian languages, etc.) and by subfamily (for example, Scandinavian languages within the Germanic language family). Generally, the classification into families and sub-families was not changed, with the exception of the following:

- (i) the families Sino-Tibetan, Tai and Austro-Asiatic were added;
- (ii) some residual categories were added to existing language families, e.g. "Germanic languages, n.i.e. (not included elsewhere)" was added to the family "Germanic languages";
- (iii) several aboriginal languages were included in 1991 in a residual category, "Athapaskan languages, n.i.e.", instead of being identified separately, due to their small size.

As a result, some individual language categories are not identical for the 1986 and 1991 Censuses. For example, the total for "Chinese" for 1986 is equivalent to the sum of the categories "Chinese" and "Sino-Tibetan languages, n.i.e." for 1991.

Overall, persons with neither English nor French as **mother tongue** were classified into more than 100 linguistic groups in 1991 compared with under 80 in 1986; a comparative description of language categories used in the 1991, 1986 and 1981 Censuses is provided in Appendix C of **The 1991 Census Dictionary**, Catalogue No. 92-301E, Statistics Canada, Ottawa.

These changes in the classification of languages, however, should not affect the comparability of data from the previous censuses for the larger language groups.

V. Edit and Imputation

Editing the collected data is one of the most important steps in the development and ultimate dissemination of the 1991 Census. This stage of census operations involves judgement as to the consistency of responses. This is to be followed by an assessment of the number of missing or invalid responses. Errors may be the result of respondents answering questions incorrectly or incompletely, or they may have been generated during coding or data capture activities. After errors are detected, values for missing or incomplete entries are imputed.

Imputation, which involves the automated assignment of responses to cases which are either invalid or missing, is of fundamental importance in following the edit procedures. The major concern in imputation is to retain as many of the original responses as possible.

Imputation may be done using either the "deterministic" or "hot deck" method - both were used in the 1991 Census. "Deterministic" imputation involves correcting errors by inferring the appropriate value on the basis of other answers on the questionnaire. This method is referred to as the "family imputation method". On the other hand, the "hot deck" approach involves selecting a record (i.e. a donor) that has a number of characteristics in common with the record in error (i.e. a recipient), and imputing the missing information from this "donor" record. Imputation is based on a consistent record, taking into account certain geographic constraints. For example, there is a high probability of imputing a missing mother tongue response as "French" in areas of the country in which the overwhelming majority of the population has French as mother tongue.

The entire edit and imputation procedure consisted of three major processes, namely the pre-derive process, imputation by the "hot deck" method, and the post-derive process. The types of responses corrected or imputed under each of these methods are as follows:

1. The pre-derive process

In this "deterministic" process, the following types of records were corrected:

- write-ins corresponding to English and/or French were provided;
- a pseudo-language was reported (i.e. a respondent who may have immigrated to Canada was reporting his country of birth as **mother tongue**, such as a person reporting "Belgian" as mother tongue rather than French or Flemish);
- "Indian", "Indien" or "Indian Dialect" was reported.

Also under this process, parents with missing data on mother tongue were assigned the language(s) of their children - if the latter had provided data.

2. Imputation by the "Hot Deck" method

Records with the following conditions were considered for this process:

- no response was provided;
- non-classifiable codes were present;
- a pseudo-code was assigned at the automated coding stage; for instance, a respondent may have entered "Canadian" or "babytalk" as mother tongue (such entries were considered "invalid" and treated as non-responses).

This process consisted of two distinct phases. In the first phase, blank responses were imputed by a probabilistic procedure known as the "hot deck" method. This technique was applied in situations where information on the mother tongue of a person was missing and a valid answer could not be found from other family members' responses. A valid "donor" record with characteristics identical to or very similar to those of the record with missing data was located using the hot deck method and the missing data were transferred to the recipient record. Each person was imputed individually. As a consequence, parents and spouses were imputed separately, without consideration of the language characteristics of other family members. Moreover, two spouses did not have to receive their imputed values from the same donor record.

The hot deck method was divided into two strata: an "aboriginal stratum" and a "non-aboriginal stratum". Donor records had to match on family status (i.e. "Husband", "Wife", "Lone Parent", or "Non-member"), and on age category ("Less than 25 years old", or "25 years or older").

In the second phase of imputation by "hot deck", missing responses were corrected by the "default" procedure. That is, when no valid donor record was found within the appropriate stratum after 250 attempts, responses were imputed randomly on the basis of the missing record's stratum and province of residence. For example, for non-matching records in British Columbia, 84% of the imputation was to English, the remaining 16% to Chinese.

3. The post-derive process

In this process, data on mother tongue were assigned for children in families on the basis of mother tongue data either of other children in the same family (where available) or of parents. In other words, there were two steps involved in this process, as shown in Figure 1.

Figure 1. Schematic Process of Family Imputation

Process Name	Action
Pre-derive: Family Imputation 1	Impute language to parents from language(s) of children, if provided.
Post-derive: Family Imputation 2.1	Impute language to children from language(s) of other children, if provided.
Family Imputation 2.2	Impute language to children from language(s) of parents.

4. Magnitude of imputation performed on data on mother tongue

Of the total population of 27,296,860, data on **mother tongue** of 542,620 persons were imputed - thus giving the national imputation rate of 2% for mother tongue. These persons had either not provided any response to the question on mother tongue or had provided some invalid or non-classifiable responses such as "babytalk" or "Canadian".

Of the entire imputation, over two-thirds was done by the "hot deck" procedure and one-fourth involved assigning the parents' language(s) to the children. Situations where mother tongue was assigned from one or more siblings to others within the same family accounted for another 6% of the total imputation performed (Table 5.1). It is interesting to note that the imputation rates in respect to mother tongue were the same for both Quebec and the rest of Canada.

Census of Population - Reference Products
1991 Census Technical Reports

Table 5.1 Imputation for Mother Tongue by Type(1), Canada, Quebec and Canada Excluding Quebec, 1991 Census

Type of imputation	Canada		Quebec		Canada excluding Quebec	
	Number	%	Number	%	Number	%
Family Imputation 1	3,985	0.7	665	0.5	3,320	0.8
Hot deck: donors' records	362,495	66.8	96,380	68.8	266,115	66.1
Hot deck: default	1,150	0.2	150	0.1	1,000	0.2
Family Imputation 2.1	29,905	5.5	5,820	4.2	24,085	6.0
Family Imputation 2.2	145,085	26.7	37,085	26.5	108,000	26.8
Total number of persons with imputation	542,620	100.0	140,100	100.0	402,520	100.0
Total number of persons	27,296,860	...	6,895,960	...	20,400,900	...
Proportion of persons with imputation (%)	2.0	...	2.0	...	2.0	...

(1) See text for explanation of types of imputation.

The imputation rate by mother tongue (for those who had provided only one response to the question on mother tongue) was the lowest (1.8%) for those reporting French and the highest (2.6%) for those reporting neither English nor French as mother tongue. Most of the imputation performed for each of these three language groups involved filling in blanks by the "hot deck" procedure. Assignment of the parents' mother tongue to the children was much more predominant among persons with English only as mother tongue than for persons with either French or a non-official language as mother tongue (Table 5.2). The mother tongue with the highest frequency of assignment was English. Of the 542,620 persons whose mother tongue was imputed, only 5,770 (1.1%) were assigned more than one mother tongue.

Table 5.2 Imputation for Mother Tongue by Type(1), for Persons With Single Response, 1991 Census

Type of imputation	English		French		Non-official language	
	Number	%	Number	%	Number	%
Family Imputation 1	2,925	0.9	520	0.4	500	0.5
Hot deck: donors' records	210,395	64.1	83,470	71.8	68,600	74.4
Hot deck: default	815	0.2	165	0.1	175	0.2
Family Imputation 2.1	16,775	5.1	3,795	3.3	7,635	8.3
Family Imputation 2.2	97,520	29.7	28,230	24.3	15,330	16.6
Total number of persons with imputation	328,430	100.0	116,180	100.0	92,240	100.0
Total number of persons	16,516,180	...	6,505,565	...	3,549,305	...
Proportion of persons with imputation (%)	2.0	...	1.8	...	2.6	...

(1) See text for explanation of types of imputation.

It may be noted from Table 5.3 that of all persons with assigned mother tongue data, 12.4% were under 5 years of age, 17.2% between 5 and 19 years, 54.8% between 20 and 64, and 15.5% were 65 years of age and over. Most of the persons under 20 years of age were assigned the mother tongue(s) of their parents (i.e. using Family Imputation 2.2) whereas mother tongues of almost all those 20 years of age and over were imputed by the "hot deck" method. Overall, of the 542,620 persons with imputed mother tongue data, 264,280 (or 48.7%) were aged 20 to 64 and had either not answered the question on mother tongue or had provided some non-classifiable response.

Table 5.3 Imputation for Mother Tongue by Type(1) and Age Groups, 1991 Census

Type of Imputation	0 - 4		5 - 19		20 - 64		65 and over		Total	
	Number	%	Number	%	Number	%	Number	%	Number	%
Family Imputation 1	0	0.0	15	0.0	3,595	1.2	370	0.4	3,965	0.7
Hot deck: donors' records	4,670	6.9	10,545	11.3	264,280	88.8	83,000	98.4	362,495	66.8
Hot deck: default	0	0.0	0	0.0	240	0.1	910	1.1	1,155	0.2
Family Imputation 2.1	18,160	26.9	9,690	10.4	2,055	0.7	5	0.0	29,900	5.5
Family Imputation 2.2	44,705	66.2	72,880	78.3	27,445	9.2	55	0.1	145,085	26.7
Total number of persons with imputation	67,530	100.0	93,125	100.0	297,610	100.0	84,350	100.0	542,620	100.0
Total number of persons	1,906,500	...	5,654,690	...	16,565,700	...	3,169,965	...	27,296,860	...
Proportion of persons with imputation (%)	3.5	...	1.6	...	1.8	...	2.7	...	2.0	...

(1) See text for explanation of types of imputation.

Although the overall imputation rate for persons with neither English nor French as mother tongue was 2.6%, the rate varied anywhere between 1% and 10% or even more, depending on their mother tongue. For example, of the 105 language groups, 14 had an imputation rate between 1.0% and 1.9% and another 77 had a rate between 2.0% and 4.9%. There were 12 language groups that had an imputation rate of 5% and more; these included Baluchi (20.0%), Algonquian languages n.i.e. (13.0%), Ojibway (10.2%), North Slave (Hare) (11.1%), Iroquoian languages n.i.e. (11.1%), Salish languages (10.9%), Wakashan languages (9.0%), Mohawk (7.6%), Montagnais-Naskapi (7.4%), Cree (5.9%), Tlingit (5.6%), and Semitic languages, n.i.e. (5.0%). Since the imputation rate for a given non-official language group is very much affected by the number of donors and recipients present in that group, estimates for groups with imputation rates in excess of 5% should be interpreted with some caution.

5. Means to check quality of data

During the processing of the data, "flags" were created to indicate the original response or the action taken during editing and imputation (e.g., original pseudo-languages, uncodeable responses and responses such as "babytalk" were identified, as well as those records for which no imputation was required). Such flags may be used to further assess the quality of data on mother tongue by areas or some other family characteristics.

VI. Data Evaluation

Throughout the census-taking process, care was taken to ensure high-quality results. However, errors can still arise at virtually any stage of the census process. Some errors occur at random and tend to cancel each other out when individual responses are aggregated to a large group. On the other hand, some errors occur more systematically and may have more serious implications on estimates than random errors.

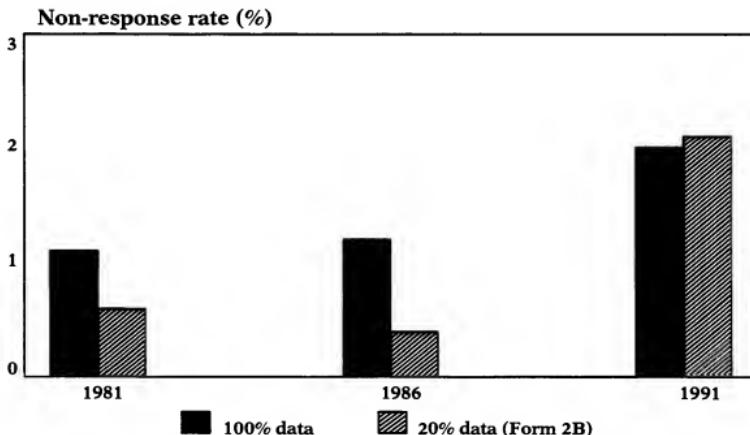
The principal types of errors that can occur in the census data are: (i) coverage errors, (ii) non-response errors, (iii) response errors, (iv) processing errors, and (v) sampling errors (applicable only to questions asked in the 2B questionnaire since it was filled out by only one-fifth of all households). **Coverage errors** occur when individuals and/or households are missed, incorrectly included, or double counted. **Non-response errors** occur when, for one reason or another, responses are not available, whereas **response errors** arise when respondents provide an incorrect response, for example, due to some misinterpretation of the wording of the question. Finally, **processing errors** can originate from data which are captured or transformed from write-in responses to numeric codes, or imputed when responses are missing or incorrectly specified.

In this section, we evaluate data in terms of **non-response errors**, **multiple responses**, **invalid responses**, and comparison of **data before and after imputation**. Although data before imputation may contain some response errors arising from invalid responses or write-ins, these should mostly reflect the genuine distribution of population by mother tongue, whereas those after imputation should reflect the impact of correcting blank or non-classifiable responses on such a distribution.

1. Non-response errors

Non-response errors are studied in terms of **non-response rate**, which is defined as the number of persons who did not answer the question on mother tongue, expressed as a percentage of all persons. Since the question on **mother tongue** was included in both the 2A and 2B questionnaires, we look at non-response rates for 100% and 20% data and their comparability by province. Again, how did the 1991 non-response rates for mother tongue differ from those observed in the 1986 and 1981 Censuses?

Figure 2. Comparison of Non-response Rates for Mother Tongue



Source: 1981, 1986 and 1991 Censuses

Table 6.1 shows that for 100% data, the non-response rate for **mother tongue** was 2.0%. The Atlantic provinces had non-response rates below the national average while provinces such as Ontario and British Columbia had rates above the national average. The variation in non-response rates among provinces may largely be attributed to the varying sizes and make-up of their linguistic communities. For example, about 99% of the population in Newfoundland and 94% in Nova Scotia had English as mother tongue, compared with 75% in Manitoba. In Ontario (where more than half of all recently arrived immigrants live) and in the four Western provinces (with a relatively greater concentration of post-war immigrants of European descent), the proportions of the population with neither English nor French as mother tongue were found to be higher than for the provinces East of Ontario.

However, it is worth noting that despite changes in the long questionnaire (Form 2B) of the 1991 Census (see section on **Concepts and Definitions**) aimed at improving data on languages (including mother tongue), the national non-response rate for mother tongue, based on the 20% sample, was slightly higher (2.1%) than that found for the 100% (or integrated) data. The differences in non-response rates based on the 100% and 20% sample data (Form 2B) were rather pronounced for Quebec, Manitoba, Saskatchewan, Alberta, and British Columbia. In fact, with the exception of Quebec, Ontario and Nova Scotia, non-response rates based on the 20% data for all provinces were higher than their counterparts from the 100% data. Overall, the range of non-response rates for mother tongue across provinces shown by 2B data amounted to 1.9 percentage points compared with 1.5 percentage points for the integrated data. Much of the difference between non-response rates for 100% and 20% data may be attributed to the underlying differences in their collection (see the last section entitled **Difference in Estimates Compiled from the 2A and 2B Forms**).

Table 6.1 Non-response Rates for Mother Tongue, Canada, Provinces and Territories, 1981, 1986 and 1991 Censuses

Province/territory	100% Data			Sample data (20%)		
	1981	1986	1991	1981	1986	1991
Percentage						
Canada	1.07	1.17	1.99	0.56	0.38	2.09
Newfoundland	0.97	0.68	0.93	0.55	0.28	0.99
Prince Edward Island	0.61	0.96	1.15	0.26	0.38	1.15
Nova Scotia	0.75	1.11	1.27	0.51	0.34	1.21
New Brunswick	0.70	0.90	1.33	0.41	0.33	1.39
Quebec	0.85	1.21	2.03	0.48	0.29	1.66
Ontario	1.01	1.15	2.28	0.52	0.39	2.26
Manitoba	1.11	1.35	1.59	0.67	0.57	2.84
Saskatchewan	0.89	0.96	1.23	0.48	0.44	2.21
Alberta	1.30	1.01	1.30	0.62	0.34	1.80
British Columbia	1.85	1.48	2.38	0.84	0.51	2.85
Yukon Territory	4.97	2.55	6.94	3.10	2.87	8.84
Northwest Territories	2.71	2.08	2.58	1.88	2.27	2.58

Irrespective of the type of data used (100% or 20%), the non-response rate for **mother tongue** was relatively higher in the 1991 Census than in the 1986 and 1981 Censuses. The 1991 non-response rate of 2.0% for mother tongue, however, should not be considered too high when compared to non-response rates of 2.3% for Marital Status and 1.9% for Sex - variables which are comparatively more simpler to respond to (a more detailed evaluation of these variables can be found in **Age, Sex, and Marital Status and Common-law Status**, Statistics Canada, Catalogue No. 92-325).

2. Multiple response

A multiple response rate is defined as the number of persons with more than one mother tongue reported, expressed as a percentage of all persons. The multiple response rate in 1991 was 2.7%, compared with 3.8% in 1986 and 2.5% in 1981. Such a variation in multiple response rates over the last three censuses can partly be attributed to the different strategies adopted for collecting data on mother tongue. For instance, in the 1981 Census, persons were to check-in only one response to mother tongue as multiple responses were not authorized, whereas in the 1986 Census, persons were authorized but not encouraged to enter more than one mother tongue. In the 1991 Census, persons were again authorized to enter more than one mother tongue but this time the question on mother tongue (in the 2B questionnaire) followed other questions on the knowledge of languages. This change in the formatting of language questions likely reduced the occurrence of a multiple response to the question on mother tongue in the 1991 Census.

Permitting respondents to enter multiple responses to the question on **mother tongue** can be beneficial as well as problematic. If on the one hand, it may encourage respondents to provide genuine data on mother tongue at a given point in time, it may also, on the other hand, encourage them to report erroneously several mother tongues. As a result, a given census may show a multiple response rate on mother tongue which is somewhat higher than the prevailing actual rate. For instance, Lachapelle (1991) has shown that the multiple response rate is about 1% or less, compared to 2.7% obtained from the 1991 Census. Again, since very few persons report consistently, from one census to another, the same two or more mother tongues (see Lachapelle, 1991), this poses another serious problem while comparing multiple response rates from different censuses.

The multiple response rate has been consistently lower for the 20% data than for the 100% data. The lowest multiple response rate for the 20% data in 1991 may be due to the **better-blocked** layout of questions on languages, which in turn may have guided respondents to enter a correct response after comprehending the difference between mother tongue, language spoken at home and knowledge of languages.

It is worth noting that regardless of the type of data, the rate of multiple response has been consistently higher in Manitoba than that in any other province of Canada. Also, the multiple response rate in Ontario has been higher than that in Quebec - the difference in rates for these two largest provinces has remained almost unchanged from 1981 to 1991.

Of the 725,805 persons who reported more than one mother tongue in the 1991 Census, 89.7% had English as one of the mother tongues compared to 40.4% of the total with French as one of the mother tongues.

Table 6.2 Multiple Response Rates for Mother Tongue, Canada, Provinces and Territories, 1981, 1986 and 1991 Censuses

Province/territory	100% Data			Sample data (20%)		
	1981	1986	1991	1981	1986	1991
Percentage						
Canada	2.46	3.75	2.66	2.24	3.43	1.29
Newfoundland	0.26	0.39	0.27	0.22	0.38	0.21
Prince Edward Island	0.72	1.30	0.84	0.65	1.23	0.26
Nova Scotia	0.89	1.40	0.89	0.82	1.30	0.38
New Brunswick	2.24	3.51	2.02	1.99	3.38	0.69
Quebec	2.41	3.67	2.59	2.12	3.42	1.08
Ontario	2.69	4.16	3.07	2.48	3.78	1.40
Manitoba	3.63	5.79	3.77	3.38	5.19	2.23
Saskatchewan	2.73	4.05	2.46	2.55	3.69	1.57
Alberta	2.46	3.75	2.63	2.34	3.37	1.58
British Columbia	2.35	3.35	2.39	2.18	3.02	1.25
Yukon Territory	2.59	2.04	1.64	2.47	1.90	1.58
Northwest Territories	3.03	3.36	2.05	3.11	3.54	2.06

3. Invalid responses

In 1991, invalid responses could be grouped into three categories:

- (i) pseudo-languages, that is responses referring to states which themselves contain more than one language community (e.g., Belgian, Scandinavian, Swiss, Czechoslovakian);
- (ii) write-ins of official languages (English and/or French);
- (iii) non-classifiable responses (e.g., "babytalk", "Canadian").

In the imputation process, different strategies were followed for these three types. Pseudo-languages were apportioned randomly to the component languages (for instance, French and Flemish in the case of a response of "Belgian"), according to predetermined distributions. Write-ins of official languages were treated as though they were checked off, while non-identifiable responses were changed to non-responses and subsequently imputed in the same manner in which other non-response items were imputed.

Of the 211,765 invalid responses, 178,023 (or 84.1%) involved write-ins of official languages and only 12,184 (or 5.8%) fell into the category of pseudo-languages (Table 6.3). Almost all of the non-classifiable responses, such as "babytalk" or "Canadian", (accounting for 21,558 or 10.2% of the total invalid responses) were treated as non-responses and imputed by the "hot deck" method.

Table 6.3 Distribution by Type of Invalid Responses to Mother Tongue, 1991 Census

Type	Number	%
Total	211,765	100.0
Pseudo-languages	12,184	5.8
Non-classifiable responses	5,275	2.5
Others(1)	16,283	7.7
Write-ins of official languages	178,023	84.1

(1) Includes "babytalk", "Canadian", etc.

Of the 178,023 persons with write-ins of official languages, 85,670 (or 48.1%) had not checked-in any box but written English and/or French as mother tongue in the blank space provided. After corrections, 71,873 of these write-ins were changed to English, 9,796 to French, and 4,001 to both English and French (Table 6.4).

Table 6.4 Distribution of Write-ins of English and/or French as Mother Tongue, as Captured and After Corrections, 1991 Census

As captured	Total	After corrections		
		English	French	English and French
Checked box:				
English	43,246	36,497	—	6,749
French	42,213	—	6,635	35,578
English and French	6,894	—	—	6,894
No box checked	85,670	71,873	9,796	4,001
Total	178,023	108,370	16,431	53,222

4. Data before and after imputation

Persons who did not respond to the question on **mother tongue** were assigned data by means of imputation – either using a “pre-derived” imputation process or other processes such as “hot deck” imputation, “default” imputation, or use of family records (see the previous section entitled **Edit and Imputation**). Out of the total population of 27,296,860 persons in the 1991 Census, data on mother tongue were imputed for 542,620 persons (consisting of those who did not respond to the question on mother tongue, or provided a non-classifiable response, or other responses such as “babytalk” or “Canadian”). Responses from 26,564,033 persons were accepted, whereas responses from 190,207 persons were considered invalid (i.e. write-ins of official languages and pseudo-languages) and thus were corrected outside the imputation process. In total, there were 26,754,240 persons (i.e. 26,564,033 + 190,207) whose data on mother tongue were not imputed.

In view of the fact that only a small proportion of persons was assigned mother tongue data, the distributions before and after imputation of population by categories of mother tongue did not show much variation. Of all persons with no imputation, 97.3% had reported English, French, or a non-official language as a single mother tongue; this proportion remained unchanged after imputation (Table 6.5).

Table 6.5 Distribution of Population by Mother Tongue, Before and After Imputation, 1991 Census (100% Data)

Mother tongue	Population before imputation		Population after imputation		Population imputed		Imputation rate (%)
	Number	%	Number	%	Number	%	
Canada							
Total	26,754,235	100.0	27,296,860	100.0	542,625	100.0	2.0
English	16,187,750	60.5	16,516,175	60.5	328,425	60.5	2.0
French	6,389,385	23.9	6,505,565	23.8	116,180	21.4	1.8
Non-official language	3,456,990	12.9	3,549,330	13.0	92,340	17.0	2.6
English and French	215,400	0.8	216,580	0.8	1,180	0.2	0.5
English and non-official language	401,215	1.5	405,080	1.5	3,865	0.7	1.0
French and non-official language	46,805	0.2	47,270	0.2	465	0.1	1.0
English, French and non-official language	29,055	0.1	29,155	0.1	100	0.0	0.3
Non-official languages	27,610	0.1	27,720	0.1	110	0.0	0.4
Quebec							
Total	6,755,865	100.0	6,895,960	100.0	140,095	100.0	2.0
English	582,740	8.6	601,410	8.7	18,670	13.3	3.1
French	5,496,460	81.4	5,597,930	81.2	101,470	72.4	1.8
Non-official language	499,590	7.4	517,985	7.5	18,395	13.1	3.6
English and French	90,930	1.3	91,590	1.3	660	0.5	0.7
English and non-official language	26,615	0.4	27,005	0.4	390	0.3	1.4
French and non-official language	38,035	0.6	38,470	0.6	435	0.3	1.1
English, French and non-official language	18,580	0.3	18,640	0.3	60	0.0	0.3
Non-official languages	2,935	0.0	2,950	0.0	15	0.0	0.5
Canada excluding Quebec							
Total	19,998,370	100.0	20,400,895	100.0	402,525	100.0	2.0
English	15,605,010	78.0	15,914,765	78.0	309,755	77.0	1.9
French	892,925	4.5	907,635	4.4	14,710	3.7	1.6
Non-official language	2,957,400	14.8	3,031,345	14.9	73,945	18.4	2.4
English and French	124,470	0.6	124,990	0.6	520	0.1	0.4
English and non-official language	374,600	1.9	378,075	1.9	3,475	0.9	0.9
French and non-official language	8,770	0.0	8,800	0.0	30	0.0	0.3
English, French and non-official language	10,475	0.1	10,515	0.1	40	0.0	0.4
Non-official languages	24,675	0.1	24,770	0.1	95	0.0	0.4

5. Difference in estimates compiled from the 2A and 2B forms

First and foremost, compared to estimates from 20% data, the 100% data showed an underestimation of 441,745 persons whose mother tongue is neither English nor French (Table 6.6). More than half of this underestimation was concentrated in Ontario - a province which has a relatively higher proportion of recently arrived immigrants.

Since four-fifths of the persons covered in the integrated data were asked to fill in a short form containing only a question on mother tongue, it is likely that persons with neither English nor French as mother tongue had (1) inadvertently reported English (or French) as their mother tongue (treating it as a principal language); and (2) provided either no response due to some confusion over the question, or reported more than one mother tongue, such as English (or French) and some other language. In consequence, compared to 20% data, integrated (100%) data overestimated by 395,555 the total number of persons with a multiple response to the question on mother tongue (Table 6.6), with Ontario accounting for 42.3% of these persons.

Data from the 2B questionnaire resulted in a relatively lower number of persons with more than one mother tongue. This is largely due to the presence of other questions on languages which, in turn, helped respondents to understand the difference between the concepts of mother tongue, language spoken at home and knowledge of languages - all put together in a **block format**.

Other factors accounting for differences in estimates from the integrated and the 20% sample data are:

- (i) estimates from the 20% sample data are subject to sampling and non-sampling errors and are based on statistical weighting procedures;
- (ii) the 20% sample data excluded persons living in institutions (such as nursing homes, hospitals and penitentiaries).

Table 6.6 Distribution of Population by Categories of Mother Tongue and Sources of Data, Canada, Provinces and Territories, 1991 Census

Province	Single response			Multiple response			Total
	English	French	Non-official language	English and French	English and non-official language	French and non-official language	
Integrated data							
Canada	16,516,180	6,505,565	3,549,305	26,571,050	216,585	405,085	47,270
Newfoundland	559,620	2,400	1,945	566,965	790	640	20
Prince Edward Island	121,765	5,415	1,485	128,670	850	205	5
Nova Scotia	838,375	34,005	19,555	891,935	5,120	2,450	120
New Brunswick	463,975	236,385	8,690	709,250	13,400	1,015	85
Quebec	601,405	5,597,930	517,975	6,717,310	91,590	27,005	38,470
Ontario	7,578,690	464,840	1,732	9,775,715	70,860	21,870	5,695
Manitoba	7,579,935	46,925	203,940	1,050,805	7,330	31,095	445
Saskatchewan	821,525	19,695	123,375	964,590	4,085	18,495	270
Alberta	2,668,650	51,100	358,735	2,478,485	11,505	9,950	1,025
British Columbia	2,606,530	45,265	551,720	3,203,520	60,610	11,115	1,515
Yukon Territory	24,480	815	2,050	27,345	125	305	5
Northwest Territories	31,230	1,380	23,855	56,470	135	985	20
Canada	16,169,875	6,502,865	3,991,050	26,663,790	91,895	186,820	22,550
Newfoundland	555,645	2,770	4,940	563,350	180	380	—
Prince Edward Island	120,590	5,590	1,560	127,735	315	40	—
Nova Scotia	830,120	24,613	887,810	9,645	1,230	105	20
New Brunswick	460,535	241,365	711,730	4,185	480	40	25
Quebec	599,145	5,565,105	580,010	6,735,260	12,395	17,235	3,590
Ontario	7,380,370	485,395	1,969,415	9,835,175	31,395	93,935	3,485
Manitoba	784,210	49,130	226,045	1,059,390	2,950	15,200	260
Saskatchewan	807,110	20,885	136,185	964,180	1,545	9,375	215
Alberta	2,931,115	53,715	401,550	2,486,375	5,350	24,030	510
British Columbia	2,545,390	48,335	614,830	3,209,160	4,625	28,690	685
Yukon Territory	24,410	885	2,070	27,340	65	220	15
Northwest Territories	31,140	1,375	23,740	56,255	130	990	20
Canada	346,305	2,700	-441,745	-92,740	124,690	218,265	24,720
Newfoundland	3,975	-370	5	3,615	610	260	20
Prince Edward Island	1,175	-175	-75	935	535	165	5
Nova Scotia	8,255	-630	-1,505	4,125	3,455	1,220	15
New Brunswick	3,440	-3,980	-3,955	-2,500	9,215	535	45
Quebec	198,320	-21,355	-62,035	-17,950	52,105	14,770	21,260
Ontario	15,725	-2,205	-22,105	-8,585	39,465	117,935	2,210
Manitoba	14,415	1,190	-12,810	-410	2,540	15,895	185
Saskatchewan	37,535	26,165	-42,815	-7,890	6,155	25,900	515
Alberta	61,040	-570	-63,110	-5,640	6,160	31,920	430
British Columbia	70	-50	-20	5	60	85	-10
Yukon Territory	90	5	115	215	5	-5	0
Northwest Territories						0	0
Canada						0	5
						0	210

VII. Historical Comparability

Since census data provide a unique opportunity to look at various socio-demographic facets of a nation, it is natural that many analysts and policy makers would like to compare data from the current and previous censuses in order to detect specific trends or changes that have occurred over time in the socio-demographic fabric of the nation. Any comparison over time of census data would not be robust unless it is accompanied by clear statements underlying changes that took place from census to census - be they related to wording of questions, instructions that accompanied questions, coverage, collection, editing and imputation procedures, etc. A meaningful comparison of data would require no change or small changes in these factors.

Are data on **mother tongue** from the 1991 Census and previous censuses comparable? The answer is "yes", provided we keep in mind the following differences underlying the compilation of these data.

1. Wording of the question and accompanying instructions

The wording of the question has changed relatively little since the 1941 Census. The alternative specifications of "first language learned in childhood" and "first language spoken in childhood" are essentially equivalent. While the question always asked for a single response, the instructions to the respondent stressed the need for a single response in the censuses up to, and including, that of 1981. For official publications from the 1981 Census, only one language was retained in the case of those who had reported more than one language. In the 1986 Census, however, instructions were modified in the accompanying guide, allowing persons to report more than one mother tongue if they had learned these languages at the same time and had spoken one as frequently as the other when they were children. The phrasing of the question is similar for the 1991 Census and the 1986 Census, except for the insertion of the expression "at home" in 1991. Furthermore, an instruction was added in 1991 following the question: "If this person no longer understands the first language learned, indicate the second language learned." In 1986, this instruction was found in the guide which accompanied the questionnaire.

There was a change in the format of the 1991 questionnaire with questions written in terms of a "third person pronoun - **This Person**", compared with the 1981 and 1986 questionnaires using a "second person pronoun - **You**". This change in the "indirect" versus "direct" manner of soliciting data may also affect the actual response entered - especially when the questionnaire is filled in by a proxy.

2. Collection strategy

A major change in overall data collection methods occurred in 1971, when census data were collected by self-enumeration. Before 1971, census data were collected by census enumerators. While this change in method reduced bias attributable to the enumerator, other factors affecting data quality arose (such as the propensity to give multiple responses to questions which were only asking for a single response).

3. Multiple responses

According to Kraft (1980), "a part of the multiple response is reflective of reality, i.e. the respondent actually has two or more languages first spoken or learned and which are still understood" (p.78). One of the factors responsible for a multiple response to the mother tongue question, according to the author, is "the misinterpretation of the question on the part of respondents" (p.77). **Multiple responses** to the mother tongue question have been observed in all Canadian censuses from 1971 onward; it is possible that earlier censuses also received multiple responses to the question, but nothing is known about their prevalence.

The reaction to such multiple responses has varied over time. Different data collection methods have been tried in order to minimize the occurrence of multiple responses. For instance, in 1971 (when the question quite clearly specified that a single response was required) multiple responses were assigned before the edit and imputation phase, namely at the data capture stage (see Demers, 1979; Demers and Kraft, 1984). In 1981, multiple responses were also edited for publication purposes, but the unedited data were preserved for more detailed analyses (which was not done for the 1976 Census data). In 1986, multiple responses were published. While the question was still phrased in the singular, the instruction to the respondent to report one language only was eliminated.

In 1991, some changes were introduced at the data collection stage to reduce, where possible, the prevalence of multiple responses. On the long form (completed by a 20% sample of private households) a new question was introduced regarding the knowledge of language(s) other than English or French. Moreover, all the language questions on the long form were grouped together, instead of following the earlier practice of asking the mother tongue question separately from the other language questions. In the 1991 long form, the question on mother tongue followed all the other questions on languages. It was hypothesized that this grouping of questions would clarify their intent for respondents, and therefore provide more accurate data. This innovation indeed had the desired effect as the 1991 data show a pronounced difference in multiple responses between the long form (Form 2B) and the short form (Form 2A), especially in comparison with the data from the preceding censuses (see Table 7.1).

Table 7.1 Prevalence of Multiple Responses to Mother Tongue Questions by Form, Canada, 1981, 1986 and 1991 Censuses

Year	Form		
	2A (80%)	2B (20%)	Integrated (100%)
%			
1981	2.5	2.2	2.5
1986	3.9	3.4	3.8
1991	3.0	1.2	2.7

4. Incomplete enumeration of Indian reserves

Another factor affecting the historical comparability of the mother tongue data involves the fact that on several Indian reserves and Indian settlements, enumeration in the 1986 Census and in the 1991 Census was either not permitted or not completed. In total, 78 geographic areas were considered as "incompletely enumerated" in 1991; these areas and their population were not included in the census totals. Because of the missing data, comparisons between 1986 and 1991 are complex and difficult. It is likely that a relatively high proportion of this population had an aboriginal language as mother tongue. Obviously, analyses of aboriginal languages will be affected by the magnitude of such non-coverage.

5. Inclusion of non-permanent residents

For the first time since 1941, both **permanent and non-permanent residents** of Canada were enumerated. Non-permanent residents are persons who hold student or employment authorizations, a Minister's permit or who are refugee claimants. These persons, numbering 223,410 (or 0.8% of Canada's total population) can be excluded from the 1991 Census in order to improve the comparability of 1991 data on mother tongue with similar data from the previous censuses conducted after 1941 (see also Table 3.1).

6. Processing, editing and imputation

The 1991 data on mother tongue, in terms of write-ins, were transformed into numeric codes using the **automated coding system**, in contrast to manual coding done in the 1986 and earlier censuses. Although the basic editing and imputation procedures did not vary between 1991, 1986 and 1981, the categories in respect to mother tongue (other than English and French) were expanded to more than 100 in 1991. This should affect the comparability of distributions by type of mother tongue for persons with neither English nor French as mother tongue. As Kraft (1980) has stated, "a 0.5% or a 1.0% increase or decrease in a specific mother tongue between one census and the next may be simply, or at least partially, a reflection of changes in processing or enumeration methodology" (p.98).

In the 1981 and 1986 Censuses, non-official languages, namely, German, Italian and Ukrainian, were specified along with a space provided for respondents to enter a non-official mother tongue. This format was discontinued in the 1991 Census. In both the 2A and 2B questionnaires used in 1991, although one empty box was provided for specifying the mother tongue other than English or French, there was, however, a provision to capture up to two languages. Such changes in collecting and capturing data on non-official mother tongues may influence the comparability from one census to another of the group with neither English nor French as mother tongue.

7. Use of 100% or 20% data for historical comparability

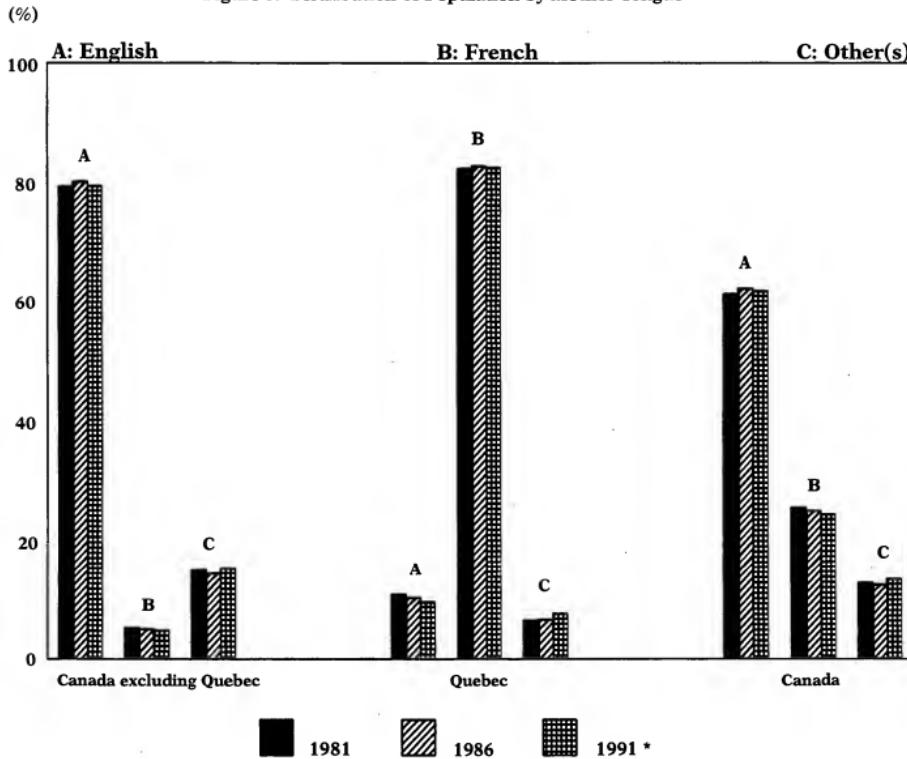
After adjusting for the change in universe (i.e. excluding non-permanent residents), 100% (or integrated) data on mother tongue from the 1991 Census can be compared with, say, data from the 1986 and 1981 Censuses. Data on language groups can be compared over time in one of the following manners:

- (i) use estimates in respect to single responses only;
- (ii) use estimates of persons having reported the language in question (single and multiple responses);
- (iii) use estimates based on the sum of single responses and some distribution of multiple responses for reported languages.

Since multiple responses vary from census to census either due to respondents' changing perceptions about their single or multiple mother tongues, to poor capability to report the same multiple mother tongues from one census to another, or to changing collection and processing procedures (Lachapelle, 1991), their presence poses a serious problem in comparing data on mother tongue from different censuses. For any meaningful historical comparisons, analysts have to handle these multiple responses by distributing them among reported language groups.

It must be noted, however, that for 1991, estimates based on the 20% sample should be used when one is interested in knowing the current number of persons by categories of mother tongue. In spite of the fact that estimates based on the 20% sample are subject to sampling errors, these estimates should reflect the more appropriate number and distribution of persons by categories of mother tongue.

Figure 3. Distribution of Population by Mother Tongue



Source: 1981, 1986 and 1991 Censuses

*Excluding non-permanent residents.

Although there are several ways in which multiple responses can be distributed (Demolinguistics Division, 1992, p.1), the simplest of these is to **distribute equally the multiple responses among reported language groups**. Based on this method, estimates of population by categories of mother tongue are compared from the 1981, 1986 and 1991 Censuses (Table 7.2).

At the national level, the proportion of the population with French as mother tongue showed a drop from 25.7% in 1981 to 24.5% in 1991. On the other hand, the number of persons with a non-official language as mother tongue has grown by 16.7% over the 1981-1991 period. These persons, in turn, accounted for 13.5% of all Canadians in 1991 compared with 12.9% in 1981.

Table 7.2 Distribution of Population by Mother Tongue, 1981, 1986 and 1991 Censuses

Mother tongue	1981		1986		1991(1)		Growth in population (%)		
	Number	%	Number	%	Number	%	1981-1986	1986-1991	1981-1991
Canada									
Total	24,343,190	100.0	25,309,330	100.0	27,073,450	100.0	4.0	7.0	11.2
English	14,684,370	60.3	15,334,085	60.6	16,463,135	60.8	4.4	7.4	12.1
French	6,127,525	25.2	6,159,750	24.3	6,498,930	24.0	0.5	5.5	6.1
Non-official	2,933,295	12.0	2,860,585	11.3	3,392,580	12.5	-2.5	18.6	15.7
English and French	208,250	0.9	332,610	1.3	216,285	0.8	59.7	-35.0	3.9
English and non-official	325,530	1.3	525,720	2.1	400,370	1.5	61.5	-23.8	23.0
French and non-official	22,255	0.1	36,315	0.1	46,200	0.2	63.2	27.2	107.6
English, French and non-official	29,475	0.1	46,585	0.2	28,975	0.1	58.0	-37.8	-1.7
Non-official languages	12,485	0.1	13,715	0.1	26,975	0.1	9.9	96.7	116.1
After adjusting for multiple responses (2)									
English	14,961,085	61.5	15,778,778	62.3	16,781,111	62.0	5.5	6.4	12.2
French	6,252,603	25.7	6,359,741	25.1	6,639,821	24.5	1.7	4.4	6.2
Other	3,129,498	12.9	3,170,826	12.5	3,652,489	13.5	1.3	15.2	16.7
Total (3)	24,343,185	100.0	25,309,345	100.0	27,073,450	100.0	4.0	7.0	11.2
Quebec									
Total	6,438,405	100.0	6,532,460	100.0	6,851,995	100.0	1.5	4.9	6.4
English	641,545	10.0	580,030	8.9	596,000	8.7	-9.6	2.8	-7.1
French	5,247,045	81.5	5,316,925	81.4	5,593,070	81.6	1.3	5.2	6.6
Non-official	394,875	6.1	393,725	6.0	486,120	7.1	-0.3	23.5	23.1
English and French	100,850	1.6	150,735	2.3	91,425	1.3	49.5	-39.3	-9.3
English and non-official	17,560	0.3	29,870	0.5	26,380	0.4	70.1	-11.7	50.2
French and non-official	18,070	0.3	30,640	0.5	37,645	0.5	69.6	22.9	108.3
English, French and non-official	17,475	0.3	29,355	0.4	18,505	0.3	68.0	-37.0	5.9
Non-official languages	980	0.0	1,185	0.0	2,850	0.0	20.9	140.5	190.8
After adjusting for multiple responses (2)									
English	706,575	11.0	680,118	10.4	661,065	9.6	-3.7	-2.8	-6.4
French	5,312,330	82.5	5,417,398	82.9	5,663,767	82.7	2.0	4.5	6.6
Other	419,495	6.5	434,950	6.7	527,145	7.7	3.7	21.2	25.7
Total (3)	6,438,400	100.0	6,532,465	100.0	6,851,995	100.0	1.5	4.9	6.4
Canada excluding Quebec									
Total	17,904,785	100.0	18,776,870	100.0	20,221,455	100.0	4.9	7.7	12.9
English	14,042,825	78.4	14,754,055	78.6	15,867,135	78.5	5.1	7.5	13.0
French	880,480	4.9	842,825	4.5	905,860	4.5	-4.3	7.5	2.9
Non-official	2,538,420	14.2	2,466,840	13.1	2,906,460	14.4	-2.8	17.8	14.5
English and French	107,400	0.6	181,875	1.0	124,860	0.6	69.3	-31.3	16.3
English and non-official	307,970	1.7	495,850	2.6	373,990	1.8	61.0	-24.6	21.4
French and non-official	4,185	0.0	5,675	0.0	8,555	0.0	35.6	50.7	104.4
English, French and non-official	12,000	0.1	17,230	0.1	10,470	0.1	43.6	-39.2	-12.8
Non-official languages	11,505	0.1	12,530	0.1	24,125	0.1	8.9	92.5	109.7
After adjusting for multiple responses (2)									
English	14,254,510	79.6	15,098,660	80.4	16,120,047	79.7	5.9	6.8	13.1
French	940,273	5.3	942,343	5.0	976,054	4.8	0.2	3.6	3.8
Other	2,710,003	15.1	2,735,876	14.6	3,125,344	15.5	1.0	14.2	15.3
Total (3)	17,904,785	100.0	18,776,880	100.0	20,221,455	100.0	4.9	7.7	12.9

(1) Excluding non-permanent residents (20% sample data).

(2) Adjustment made on the assumption that multiple responses are distributed equally among reported languages.

(3) The sum of the categories may vary slightly from the total because of rounding.

Source: Demolinguistics Division (1992), "Population by Mother Tongue, 1981 to 1991: Basic Data and Approximate Comparative Data", only for 1981 and 1986 Censuses.

VIII. Products and Services

Consultation on user needs

Greater emphasis was placed on user consultation for the 1991 Census products and services. Over the course of two years, over 3,000 organizations from the private and public sectors were approached to solicit their comments for the proposed product and service line. Solely between November 1 and November 15, 1990, eight focus groups on census data support information were surveyed for their comments and recommendations regarding the **1991 Census Technical Reports Series**. Suggested fundamental changes and improvements to the product helped meet the needs of current and potential users.

Consultation proved to be an essential exercise in developing the shape and content of the census product and service line, and in determining market potential and pricing.

Marketing of products and services

The 1991 Census Marketing Program ensures that potential data users receive the information they need on census products and services in order to make informed decisions. It seeks to reach those individuals or enterprises that rely on census data to inform them of the products and services available from the census database and their potential uses and applications. The national headquarters in Ottawa and the regional reference centres across the country work in partnership to ensure that the largest number of people possible are aware of what the census database has to offer. Therefore, with each data release, the 1991 Census Marketing Program ensures that the information relative to the release is available to the general public through many outlets, especially the media.

Products based on census data on languages

Under **The Nation Series**, basic data on each of the variables covered in the 1991 Census are provided at the national and provincial levels. Cross-tabulations by socio-demographic characteristics presented in this series may offer users the potential of more detailed analyses and/or enable them to compare current data with similar data from earlier censuses (keeping in mind the changes in coverage, collection, and other methodological procedures). In this series, data from the 1991 Census on **mother tongue** and **other language variables** have been published by Statistics Canada in the following reports:

- **Mother Tongue**, Catalogue No. 93-313.
- **Mother Tongue: 20% Sample Data**, Catalogue No. 93-333.
- **Home Language and Mother Tongue**, Catalogue No. 93-317.
- **Knowledge of Languages**, Catalogue No. 93-318.
- **Population Estimates, First Official Language Spoken, 1991**, Catalogue No. 94-320.

Data on **languages** published in these reports are further supplemented by **Basic Summary Tabulations** – tabulations featuring two or more interrelated variables (such as mother tongue and home language, or mother tongue and knowledge of official language(s)) – as well as by data published under the **Dimensions Series**. The latter examines themes of considerable public interest and historical trends. Some of the titles planned under the **Dimensions Series** are "Profile of Language Groups, 1991", and "Language Retention and Transfer, 1991". For more details about these publications, see the **1991 Census Catalogue**, Catalogue No. 92-302.

Small area data on languages or on other census variables not available in publications may be obtained by contacting the **Statistical Reference Centre** of Statistics Canada (613-951-8116).

IX. Conclusion

In this report, we have examined the quality of the 1991 Census data on mother tongue. Indicators used to examine quality of data included the phrasing of the question, coverage and collection, data assimilation, edit and imputation, data evaluation, and historical comparability with data from previous censuses. The analysis has shown that, overall, the mother tongue data collected in the 1991 Census are valid and reliable.

As in all the censuses since 1971, the mother tongue question was asked of the entire population in 1991, although the manner in which it was asked and processed changed somewhat. The non-response rate for mother tongue was 2.0% for integrated data compared to 2.1% for 20% sample data. On the other hand, the better structuring of questions on languages and their placement in a **block format** for the 20% sample data resulted in a much lower multiple response rate (1.2% compared with 2.7% for integrated data) for mother tongue. Compared with the 1986 Census, the 1991 Census had a higher non-response rate but a lower multiple response rate (for both the integrated and the 20% data).

Of the total population of 27,296,860 persons, 97.3% had provided an acceptable response to the question on mother tongue. There were 542,620 persons (or 2.0%) whose mother tongue was imputed (including those who had entered either a non-classifiable response or an invalid response such as "babytalk" or "Canadian"). Another 12,184 had entered pseudo-languages and 178,023 had "write-ins" of official languages. Responses to these two categories were corrected during the editing phase (which preceded the imputation phase). Of all those whose mother tongue was imputed, 54.8% were in the 20 to 64 age group.

For the first time since 1941, the 1991 Census included non-permanent residents in its coverage. Since these persons (most with neither English nor French as mother tongue) represent only 0.8% of the total population, their inclusion or exclusion would not have any noticeable impact on the overall distributions of population by mother tongue.

Bibliography

Ciok, Rick (1991) "The Use of Automated Coding in the 1991 Canadian Census of Population", paper presented at the 1991 Annual Meeting of the American Statistical Association, Atlanta.

Demers, Linda (1979) "Evaluation de la qualité des informations ethniques et linguistiques fournies par les recensements canadiens, 1901 à 1976", (unpublished M.A. thesis, Demography), Montréal: Université de Montréal.

Demers, Linda and Kralt, John (1984) "On the Comparability of the Census Mother Tongue Data", Technical Working Paper, Ottawa: Statistics Canada/Secretary of State.

Demolinguistics Division (1992) "Population by Mother Tongue, 1981 to 1991: Basic Data and Approximate Comparative Data", Working Paper No. 2, Ottawa: Statistics Canada.

Lachapelle, Réjean (1991) "The Use of Census Data for the Implementation of Language Legislation", Demolinguistics Division, Working Paper No. 1, Ottawa: Statistics Canada.

Kralt, J.M. (1980) "A User's Guide to 1976 Census Data on Mother Tongue" Working Paper No. 3 - DSC 79, Ottawa: Statistics Canada.

Statistics Canada: The 1991 Census Dictionary Catalogue No. 92-301E, Ottawa.

United Nations (1969), Principles and Recommendations for the 1970 Population Censuses, Statistical Papers, Series M, No. 44, Department of Economic and Social Affairs, Statistical Office of the United Nations, New York: United Nations.

Other Census Reference Products

Dictionaries and Handbooks

Statistics Canada. 1991 Census Dictionary. Ottawa: Supply and Services Canada, 1992. 1991 Census of Canada. Catalogue No. 92-301E.

Statistics Canada. 1986 Census Dictionary. Ottawa: Supply and Services Canada, 1987. 1986 Census of Canada. Catalogue No. 99-101E.

Statistics Canada. 1991 Census Handbook. Ottawa: Supply and Services Canada. 1991 Census of Canada. Catalogue No. 92-305E.

Statistics Canada. 1986 Census Handbook. Ottawa: Supply and Services Canada. 1986 Census of Canada. Catalogue No. 99-104E.

Statistics Canada. C91 Software User Guide. Ottawa: Supply and Services Canada, 1992. 1991 Census of Canada.

Publications on Language Variables

Data Products - 1991 Census

- 93-313 Mother Tongue
The Nation
- 93-333 Mother Tongue: 20% Sample Data, 1991
The Nation
- 93-317 Home Language and Mother Tongue
The Nation
- 93-318 Knowledge of Languages
The Nation
- 94-318 Profile of Language Groups, 1991 (forthcoming)
Dimensions
- 94-319 Language Retention and Transfer, 1991 (forthcoming)
Dimensions
- 94-320 Population Estimates by First Official Language Spoken
Dimensions

Reference Products - 1991 Census

- 92-336E Home Language and Knowledge of Languages (forthcoming)
1991 Census Technical Reports

Census of Population – Reference Products
1991 Census Technical Reports

Equivalent Publications from Previous Censuses

- 93-102 Language: Part 1 (1986)
The Nation
- 93-103 Language: Part 2 (1986)
The Nation
- 93-153 Language Retention and Transfer (1986)
Dimensions
- 92-902 Mother Tongue (1981)
National
- 92-910 Mother Tongue, Official Language and Home Language (1981)
National

Please note that not all publications from previous censuses are available for purchase. Copies may be referenced through the Regional Reference Centres or Depository Libraries.

Regional Reference Centres

Statistics Canada's Regional Reference Centres provide a full range of census products and services. Each reference centre is equipped with a library and a sales counter where users can consult or purchase publications, microcomputer diskettes, microfiche, maps and more.

The staff of the Regional Reference Centres provides consultative and research services in addition to providing after-sales service and support, including seminars and workshops on the use of Statistics Canada information.

Each centre has facilities to retrieve information from Statistics Canada's computerized data retrieval systems CANSIM and E-STAT. A telephone inquiry service is also available with toll-free numbers for regional users outside local calling areas. Call, write, fax or visit the nearest Regional Reference Centre for more information.

Atlantic Region

Advisory Services
Statistics Canada
Viking Building, 3rd Floor
Crosbie Road
St. John's, Newfoundland
A1B 3P2

Advisory Services
Statistics Canada
North American Life Centre
1770 Market Street
Halifax, Nova Scotia
B3J 3M3

Local calls: (902) 426-5331
Toll-free service: 1-800-565-7192
Fax number: (902) 426-9538

Quebec Region

Advisory Services
Statistics Canada
200 René Lévesque Blvd. W.
Guy Favreau Complex
Suite 412, East Tower
Montréal, Québec
H2Z 1X4

Local calls: (514) 283-5725
Toll-free service: 1-800-361-2831
Fax number: (514) 283-9350

National Capital Region

Statistical Reference Centre (NCR)
Statistics Canada
R.H. Coats Building Lobby
Tunney's Pasture
Ottawa, Ontario
K1A 0T6

Local calls: (613) 951-8116
If outside the local calling area,
please
dial the toll-free number for
your province.
Fax number: (613) 951-0581

Ontario Region

Advisory Services
Statistics Canada
Arthur Meighen Building, 10th
Floor
25 St. Clair Avenue East
Toronto, Ontario
M4T 1M4

Local calls: (416) 973-6586
Toll-free service: 1-800-263-1136
Fax number: (416) 973-7475

Prairie Region

Advisory Services
Statistics Canada
MacDonald Building, Suite 300
344 Edmonton Street
Winnipeg, Manitoba
R3B 3L9

Local calls: (204) 983-4020
Toll-free service: 1-800-542-3404
Fax number: (204) 983-7543

Advisory Services
Statistics Canada
Avord Tower, 9th Floor
2002 Victoria Avenue
Regina, Saskatchewan
S4P 0R7

Local calls: (306) 780-5405
Toll-free service: 1-800-667-7164
Fax number: (306) 780-5403

Prairie Region (Concluded)

Advisory Services
Statistics Canada
Park Square, 8th Floor
10001 Bellamy Hill
Edmonton, Alberta
T5J 3B6

Local calls: (403) 495-3027
Toll-free service: 1-800-282-3907
N.W.T.- Call collect: 1-403-495-3028
Fax number: (403) 495-5318

Advisory Services
Statistics Canada
First Street Plaza, Room 401
138-4th Avenue South East
Calgary, Alberta
T2G 4Z6

Local calls: (403) 292-6717
Toll-free service: 1-800-472-9708
Fax number: (403) 292-4958

Pacific Region

Advisory Services
Statistics Canada
Sinclair Centre, Suite 440F
757 West Hastings Street
Vancouver, British Columbia
V6C 3C9

Local calls: (604) 666-3691
Toll-free service: 1-800-663-1551
(except Atlin, B.C.)
Yukon and Atlin, B.C.
Zenith 08913
Fax number: (604) 666-4863

Telecommunications Device Access for the Hearing-impaired 1-800-363-7629

Toll free order only line (Canada and USA) 1-800-267-6677

Depository Libraries

The Statistics Canada Library in Ottawa maintains complete current and historical records of all Statistics Canada publications, both catalogued and non-catalogued. The library staff is available to help users find the required information.

Statistics Canada Library
R.H. Coats Building, 2nd Floor
Tunney's Pasture
Ottawa, Ontario
K1A 0T6
Local calls: 613-951-8219/20
Fax: 1-613-951-0939

The following is a list of full depository libraries that receive all Statistics Canada publications and all other federal government publications.

Canada

Newfoundland

St. John's
Memorial University of
Newfoundland
Queen Elizabeth II Library
St. John's, Newfoundland
A1B 3Y1

Prince Edward Island

Charlottetown
Government Services Library
Charlottetown, Prince Edward
Island
C1A 3T2

Nova Scotia

Halifax
Dalhousie University
Killam Memorial Library
Halifax, Nova Scotia
B3H 4H8

Wolfville

Acadia University
Vaughan Memorial Library
Wolfville, Nova Scotia
B0P 1X0

New Brunswick

Fredericton

Legislative Library
Fredericton, New Brunswick
E3B 5H1
University of New Brunswick
Harriet Irving Library
Fredericton, New Brunswick
E3B 5H5

Moncton

Université de Moncton
Bibliothèque Champlain
Moncton, New Brunswick
E1A 3E9

Sackville

Mount Allison University
Ralph Pickard Bell Library
Sackville, New Brunswick
E0A 3C0

Quebec

Montréal

Municipal Library of Montréal
Montréal, Quebec
H2L 1L9
Services documentaires multimédia
Montréal, Quebec
H2C 1T1

Concordia University Library

Montréal, Quebec
H3G 1M8

McGill University

McLennan Library
Montréal, Quebec
H3A 1Y1

Université de Montréal
Bibliothèque des sciences
humaines
et sociales

Montréal, Quebec
H3C 3T2

Université du Québec à Montréal
Bibliothèque
Montréal, Quebec
H2L 4S6

Québec

National Assembly Library
Québec, Quebec
G1A 1A5

Sherbrooke

Université de Sherbrooke
Bibliothèque générale
Cité universitaire
Sherbrooke, Quebec
J1K 2R1

Sainte-Foy Université Laval Bibliothèque générale Sainte-Foy, Québec G1K 7P4	University of Ottawa Morisset Library Ottawa, Ontario K1N 9A5	The University of Manitoba Elizabeth Dafoe Library Winnipeg, Manitoba R3T 2N2
Ontario	Sudbury	Saskatchewan
Downsview York University Scott Library Downsview, Ontario M3J 2R6	Laurentian University of Sudbury Library Sudbury, Ontario P3C 2C6	Regina Legislative Library Regina, Saskatchewan S4S 0B3
Guelph University of Guelph Library Guelph, Ontario N1G 2W1	Thunder Bay Lakehead University Chancellor Paterson Library Thunder Bay, Ontario P7B 5E1	Saskatoon University of Saskatchewan The Main Library Saskatoon, Saskatchewan S7N 0W0
Hamilton Hamilton Public Library Hamilton, Ontario L8R 3K1	Toronto Legislative Library Toronto, Ontario M5S 1A5	Alberta
McMaster University Mills Memorial Library Hamilton, Ontario L8S 4L6	Metropolitan Toronto Reference Library Toronto, Ontario M4W 2G8	Calgary The University of Calgary MacKimmie Library Calgary, Alberta T2N 1N4
Kingston Queen's University at Kingston Douglas Library Kingston, Ontario K7L 3N6	University of Toronto Robarts Library Toronto, Ontario M5S 1A5	Edmonton Edmonton Public Library Edmonton, Alberta T5J 2V4
London The University of Western Ontario D.B. Weldon Library London, Ontario N6A 3K7	Waterloo University of Waterloo Dana Porter Arts Library Waterloo, Ontario N2L 3G1	Legislative Library Edmonton, Alberta T5K 2B6
Ottawa Library of Parliament Canadian Government Information Section Ottawa, Ontario K1A 0A9	Windsor Windsor Public Library Windsor, Ontario N9A 4M9	The University of Alberta Library Edmonton, Alberta T6G 2J8
National Library of Canada Ottawa, Ontario K1A 0N4	Manitoba	British Columbia
	Winnipeg Legislative Library Winnipeg, Manitoba R3C 0V8	Burnaby Simon Fraser University Library Burnaby, British Columbia V5A 1S6
		Vancouver The University of British Columbia Library Vancouver, British Columbia V6T 1Y3
		Vancouver Public Library Vancouver, British Columbia V6Z 1X5

**Census of Population - Reference Products
1991 Census Technical Reports**

Victoria

Legislative Library
Victoria, British Columbia
V8V 1X4

University of Victoria
McPherson Library
Victoria, British Columbia
V8W 3H5

Northwest Territories

Yellowknife

Northwest Territories
Government Library
Yellowknife, Northwest Territories
X0E 1H0

Other Countries

Federal Republic of Germany

Preussischer Kulturbesitz
Staatsbibliothek
Abt. Amtsdruckchriften U. Tausch
Postfach 1407
1000 Berlin 30
Germany

United Kingdom

The British Library
London, WC1B 3DG
England, United Kingdom

Japan

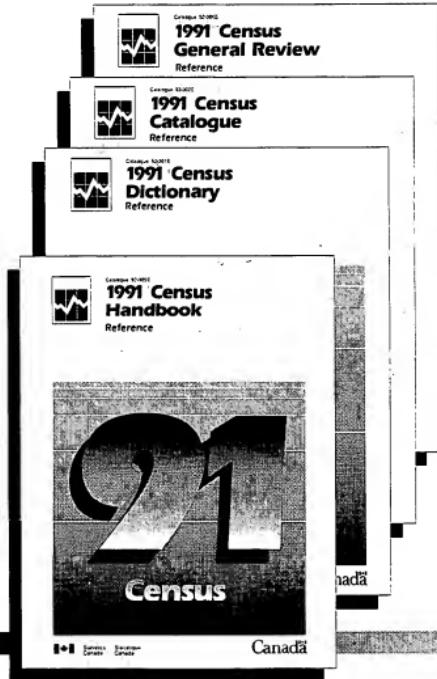
National Diet Library
Tokyo, Japan

United States of America

Library of Congress
Washington, D.C. 20540
United States of America

The Census is the most comprehensive database available in Canada today...

are you sure
you're getting
everything
out of it?



Taking full advantage of Canada's largest, most comprehensive social and economic database is often overwhelming, but the Census can be the most valuable business tool you will ever use. Statistics Canada has designed a series of 1991 General Reference Products to put the Census to work for you.

To order the *Census Dictionary*, *Census Handbook*, *Census General Review* or a *Census Catalogue* of products and services, call your nearest Statistics Canada Regional Reference Centre or our national order line...

1 800 267-6677

Nuts and Bolts of Census Data

1991 *Census Technical Reports* provide users with data quality information. Census concepts, variables and their components, definitions, coverage, processing, data evaluation and limitations and much more are explained in detail in each report.



For a complete list of *Technical Reports* available, call your nearest Statistics Canada Regional Reference Centre or our national order line...

1 800 267-6677